REMARKS

Reconsideration of this application is requested in view of the amendments to the specification and claims and the remarks presented herein.

The claims remaining in the application are claims 1 to 8 and 29 to 32 which are the elected claims. The non-elected claims have been cancelled but Applicants reserve the right to file a divisional application directed thereto. It should be noted that claim 29 should be combined with the elected invention as set forth in the restriction requirement in the office action of August 26, 2003. This is even more so since claims 30 to 32 are dependent thereupon.

Applicants have provided a new Abstract of the Disclosure and the specification has been amended to delete the embedded hyperlink codes therein. The specification has also been amended to recite the address of the Depository and the date and name of the Depository are already in the specification.

Claims 1, 2 and 30 to 32 were rejected under 35 USC 101 because the Examiner was of the opinion that the claimed invention was not supported by a specific and substantial asserted utility or a well established utility. The Examiner states that the specification discloses the open reading frame of SEQ ID No: 2, gene CaNL256 from

Candida Albicans and that the application does not give any evidence or guidance concerning the activity of the gene or if that expression or lacking of expression associated with the essential functions of the fungus. The Examiner is of the opinion that the Applicant has not adequately described any specific activity for the gene encoded by the SEQ ID No: 2. Therefore, it is doubtful whether the nucleotide sequences can have any of Applicants' asserted utilities according to the Examiner.

Applicants respectfully traverse this ground of rejection since it is deemed that the specification complies with the 35 USC 101 utility requirement. Applicants have cloned the *C.Albicans* homolog of YNL256 and it is an identification of the clone DNA is described by a homology research in the description on page 30.

With respect to the Examiner's allegation that the level of homology between YNL256 and cNL256 based on his own BLAST search. Applicants submit that the results of the BLAST search as shown in the search enclosed by the Examiner have not been properly interpreted. First, the matched DNA sequence YNL256w shown in the Examiner's enclosed documents is longer than its own open reading frame by 414 base pairs (as evidenced by the SwissProt protein sequence). The Examiner is comparing a sequence containing untranslated elements with a shorter genomic sequence isolated by the Applicant.

The matched sequence being longer than the reference, it is thus artificially reducing the numerical result "QUERY Match". As a cursory example, the 414 base pair difference in the BLAST search would lower the match percentage by 414/2364 = 18%. In addition, it is indicated in line 15 of page 30 that the coding region of CaNL256 is shorter than that of its YNL256 homolog and this would artificially lower the overall match score. Finally, the Applicant has stated in lines 17 to 18 of page 30 that account must be taken of the translational variant in *C. Albicans.*

If the BLAST search was performed without taking these into account, then the overall match score will be artificially lowered. For example, it is noticed that on the search extract submitted by the Examiner, result 7 appears to be a match between the sequence in issue and other sequence corresponding to the same gene disclosed by the same Applicant in corresponding foreign patents. Interestingly, the overall match score of the sequence itself is only 16.6% with a 99% local similarity, probably due to the size differences. Thus, the proper measure of similarity is the best local similarity which in the Examiner's search, appears to be already 49.7% with YNL256. The artificially low number of 12% query match is to be disregarded as numerical artefact produced by differing sizes. There is no reason to doubt the accuracy of the homology search performed by the Applicant and which yields similar pertinent results to that of the Examiner.

With respect to the doubts raised by the Examiner about the accuracy of the determination of the function of the cloned gene based upon the proposition that the similarity with both the DHPS and HPPK domains is unlikely, Applicants submit two publications herewith, namely, Ogata et al, Science, Vol. 293, pages 2093-2098 and Mouillon et al, Biochem. J., Vol. 363, pages 313-319. The Examiner's attention is directed to column 2 of page 2096 of Ogata et al and the Abstract of the Mouillon et al references which show that both in the plant and the bacterial kingdom, there are enzymes that have both DHPS and HPPK functions within folate and folic acid pathways. It appears to be a feature of such enzymes and therefore, it is likely that CaNL256 has all of these activities and there is no reason to doubt the Applicants' results owing to the presence of these two domains.

Finally, the homology of 52% at the nucleotide level is high enough in addition to the matching of specific highly conserved domains for a function to be assigned. It is further submitted that for unicellular organism, the mutation rate being high, there is considerable genetic drift across DNA regions which do not code for specific catalytic domains. Thus, homology results across long regions are to be lower than that in other life forms as indicated in lines 20 to 25 of page 13.

The quoted articles on functions assigned by homology do not make definite technical statements but, rather, give opinions. The Attwood et al reference states that it

is unadvisable to make functional assignments merely on "some degree of similarity". All that this could indicate is that if a homology is too low, then it would be less reliable. Moreover, with respect to Brenner, this states that without laboratory experiments, it is not possible to "know for certain". An examination of the USPTO revised guidelines shows that what is required is a credible utility not an absolute scientific certainty.

Credible utility is defined in the utility guidelines (page 5) that "Where an Applicant has specifically asserted that an invention has a particular utility, that assertion cannot simply be dismissed by Office personnel as being "wrong". Rather, Office personnel must determine if the assertion of utility is credible (i.e. whether the assertion of utility is believable to a person of ordinary skill in the art based on the totality of evidence and reasoning provided). As assertion is credible unless (A) the logic underlying the assertion is seriously flawed, or (B) the facts upon which the assertion is based are inconsistent with the logic underlying the assertion." (emphasis added) Therefore, it is deemed that claims 1, 2 and 29 to 32 comply with 35 USC 101 and withdrawal of this ground of rejection is requested.

The elected claims were rejected under 35 USC 112, first paragraph, as not being based upon an enabling disclosure since, in the Examiner's opinion, the claims contain subject matter which is not described in the specification so as to reasonably convey to one skilled in the art that Applicant had possession of the claimed invention at the time of filing the application. The Examiner states that the claims are directed to polynucleotides

of SEQ ID No: 2, homologs and functional fragments thereof are genus claims that encompass a wide variety of molecules which are not sufficiently described in the specification.

Applicants respectfully traverse these grounds of rejection since it is deemed that the homologs and functional fragments thereof are within the reach of one skilled in the art. The description teaches that the gene claimed is the homolog of YNL256w as indicated in Example 1 on pages 29 and 30. Thus, its function is that of its homologs in *S.Cerevisia* and specifically has HPPK and DHPS activity as indicated on page 31. In addition, the specification makes it clear that this gene is selected because it is essential meaning that a deletion of it is lethal as indicated in lines 9 to 11 of page 6 which can be assessed easily. Thus, one skilled in the art is perfectly able to clone by routine means, homologs of SEQ ID No: 2 and to obtain fragments thereof and select a clone suitable for working the invention. The claimed sequences can thus be obtained by one skilled in the art and are enabled.

The written requirement of 35 USC 112 does not require that <u>all</u> possible embodiments be expressly described in terms of sequence or structure to show possession of the invention. The Examiner's attention is directed to two recent decisions, namely, Enzo Biochem v. Gen-Probe Inc., 296 F.3d 1316, and Moba v. Diamond Automation, 325 F. 3d, 1306. In the Enzo case, sequences which were not sequenced or disclosed were claimed in accordance with 35 USC 112 because they were obtainable by one

skilled in the art and thus, enabled. More specifically, the Moba case held "The language of 112 indicates that a patent will contain an adequate description if it provides enough information to enable a person skilled in the art to make and use the invention."

"Any disclosure that enables one to make and use the invention also, by definition, shows that the inventor was in possession of that full invention."

"After all, to enable is to show possession and to show possession is to enable."

Clearly judicial opinion is adverse to a rule in which all sequences must be disclosed to be validly claimed especially, when they are properly enabled and as Judge Rader pointed out in the Moba case, this is impractical. In summary, the Applicant has provided structural information in the form of sequence data and functional information which would enable one skilled in the art to determine if a claimed variant is suitable for the invention and therefore, the claims comply with 35 USC 112, first paragraph, and withdrawal of this ground of rejection is requested.

Claims 1, 2 and presumably, 29 to 32 were rejected under 35 USC 112, second paragraph, as being indefinite since, in the Examiner's opinion, there was a logical disconnect between claims 1 and 2 and claims 29 to 30. The Examiner states that claim 1 recites a product and claims 29 to 32 recite a product in the preamble but have method steps in the body of the claims referring to steps that are no where in claim 1. Therefore, the inconsistency requires correction.

Applicants respectfully traverse this ground of rejection since it is deemed that Applicants are entitled to claim the invention as they see fit and claims 1 and 2 are product claims and claims 29 to 32 are product by process claims. The latter are directed to a sequence that can be obtained by a defined process and such wording is expressly permitted by the Examination Guidelines, see 2173.05(p). Claims directed to product-by-process or produce and process claims and is in the U.S. P.T.O. written description guidelines, see page 40, example 10, cited as means of overcoming the written description issues for claims to DNA sequences. These claims are clear and are proper under the guidelines. Therefore, withdrawal of this ground of rejection is requested.

All of the elected claims were rejected under 35 USC 102 as being anticipated by Accession No. SCYNL256w which, according to the Examiner, is drawn to a sequence having homology and encoding fragments of SEQ ID No: 2. The Examiner deems that the prior art discloses a product which reasonably appears to be identical or with or only slightly different from Applicants' product claims.

Applicants respectfully this ground of rejection since the reference cited by the Examiner in no way anticipates or renders obvious Applicants' invention. SCYNL256w is the homolog of CaNL256 and *S.Cerevisia* and does not have the sequence of SEQ ID No: 2 as assumed by the Examiner. The homology being 50% according to Applicants and less according to the Examiner himself and most importantly, it is not a *C.Albicans* gene and it is clear that all of the claims are properly directed to *C.Albicans* genes and

there is no reason to consider that the sequences claimed by claims 29 to 32 which are functional genes from the *C.Albicans* could have the same sequence as the sequence of SCYNL256w. The CTG code difference in line 17 of page 30 and the size differs between the known homologs makes it extremely unlikely that there could be full sequence identity. However, the trilateral project 24.1 (Biotechnology Comparative Study) indicates that for the U.S.P.T.O., it is necessary to have sequence identity for anticipation to occur and for the purposes of novelty, the prior art is neither identical nor slightly different as the Examiner suggests and therefore, the Examiner has to show that the prior art discloses an identical sequence for an anticipation rejection. Therefore, withdrawal of this ground of rejection is requested.

In view of the amendments to the specification and claims and the above remarks, it is believed that the claims clearly point out Applicants' patentable contribution and favorable reconsideration of the application is requested.

Respectfully submitted,
Muserlian, Lucas and Mercanti

Charles A. Muserlian, 19,683
Attorney for Applicants
Tel.# (212) 661-8000

CAM:ds
Enclosures

REPORTS

Mechanisms of Evolution in *Rickettsia conorii* and *R. prowazekii*

Hiroyuki Ogata,[1] Stéphane Audic,[1] Patricia Renesto-Audiffren,[2] Pierre-Edouard Fournier,[2] Valérie Barbe,[3] Delphine Samson,[3] Véronique Roux,[2] Pascale Cossart,[4] Jean Weissenbach,[3] Jean-Michel Claverie,[1]* Didier Raoult[2]*

Rickettsia conorii is an obligate intracellular bacterium that causes Mediterranean spotted fever in humans. We determined the 1,268,755-nucleotide complete genome sequence of *R. conorii*, containing 1374 open reading frames. This genome exhibits 804 of the 834 genes of the previously determined *R. prowazekii* genome plus 552 supplementary open reading frames and a 10-fold increase in the number of repetitive elements. Despite these differences, the two genomes exhibit a nearly perfect colinearity that allowed the clear identification of different stages of gene alterations with gene remnants and 37 genes split in 105 fragments, of which 59 are transcribed. A 38-kilobase sequence inversion was dated shortly after the divergence of the genus.

Rickettsia species live in different ecological niches inside different arthropod hosts (insects or ticks), in which most of them are transmitted vertically from the mother to the progeny (*1*). *R. conorii* naturally infects the dog brown tick *Rhipicephalus sanguineus*. When transmitted to humans through tick bites, the bacterium causes Mediterranean spotted fever (*1*, *2*). *R. conorii* is closely related to the previously sequenced *R. prowazekii* (*3*), the agent of louse-borne typhus. We determined the complete sequence of the *R. conorii* genome (GenBank accession number AE006914) (Table 1) (*4*, *5*). Comparative analysis of these two closely related *Rickettsia* sp. (Table 2) provides snapshots of the progression of the gene degradation process, which has been linked to adaptation to intracellular parasitism (*3*, *6*–*9*).

The overall gene order in the *R. conorii* genome [Figs. 1 and 2 and Web fig. 1 (*10*)] is remarkably similar to that of *R. prowazekii*, except for the translocation/inversion of a few short segments in the region corresponding to the end of replication. The detailed sequence comparison of the two genome sequences revealed numerous cases of apparently orthologous pairs of open reading frames (ORFs) (according to the best reciprocal match criteria) exhibiting large differences in sizes. For instance, the *phoR* gene encoding a 521-residue protein in *R. prowazekii* becomes a set of three consecutive ORFs (RC0702 through RC0704) of 643, 132, and 82 residues in *R. conorii*. Similarly, the *sca1* gene encoding a 1902-residue protein in *R. conorii* corresponds to three consecutive ORFs (RP016 through RP018) in *R. prowazekii* (Fig. 3). In addition to this "gene splitting" phenomenon (*11*), we also identified many additional *R. conorii* genes (229) and fewer *R. prowazekii* genes (6), for which a residual similarity could be found in homologous but noncoding regions in the other genome (*12*), thus representing cases of "decaying orthologs" (Fig. 2). Finally, 323 *R. conorii* ORFs exhibited no orthologous relationship (regular, split, or decaying), whereas 24 *R. prowazekii* ORFs had no equivalent in *R. conorii*. In total, 552 *R. conorii* genes have no orthologous functional counterpart in *R. prowazekii*, whereas 30 *R. prowazekii* genes have no counterpart in *R. conorii* (Table 2). Those genes are most likely to be responsible for the phenotypic differences between the two species.

Out of the 552 ORFs (reduced to 514 when every set of multiple ORFs arising from split genes is counted as one) constituting the gene excess in *R. conorii*, 106 ORFs (reduced to 79 as before) were assigned a putative function. These supplementary genes are overrepresented ($P < 0.05$, χ^2 test) as compared with those of *R. prowazekii* in the categories of DNA replication, transporters, regulatory functions, and drug sensitivity (Table 2). *R. conorii* has three genes (RC0843-4, RC0017, and RC0450) related to DNA transformation functions, including the DNA uptake protein ComF, the competence operon protein ComE3, and the chromosomal transformation protein Smf. The presence of such a DNA transformation gene has not been previously described for other obligate intracellular parasites [*R. prowazekii* (*3*), *Chlamydia* spp. (*13*), and *Mycobacterium leprae* (*9*)]. *R. conorii*'s capability for exogenous DNA uptake is further suggested by the presence of four ORFs of apparently foreign origin: one phage-related protein (RC0490), one insertion element (RC0688), and two lysozymelike proteins from viruses (RC0727 and RC1298). Both *Rickettsia* species are naturally resistant to penicillin and aminoglycoside antibiotics, and *R. conorii* exhibits higher resistance than *R. prowazekii* to antibiotics (*14*). Consistently, its genome contains nine additional genes related to its sensitivity to drugs, including four genes for β-lactamases (RC1243-4) and its regulation (RC0535, RC0788, and RC1358); three drug efflux transporters (RC0301, RC0564-9, and RC1181); an aminoglycoside 3′-phosphotransferase (RC0947); and an acetyltransferase (RC0554). *R. conorii* is known to move around inside host cells by propulsion produced by continuous actin polymerization (*15*). No clear homolog of proteins known to be responsible for the actin-based motility of *Listeria monocytogenes* (ActA) or *Shigella flexneri* (IcsA) (*16*) was found, but an ORF (RC0909) coding for a 520-residue–long protein exhibits an overall organization similar to that of ActA. Both proteins share a highly charged NH_2-terminus (~300 residues) and a central proline-rich region. RC0909 has a weak similarity to the WASP homology domain 2, found in a family of proteins known to regulate the formation of the actin filaments.

As intracellular parasites, *Rickettsia* have small genomes and an evolutionary tendency toward further genomic reduction (*6*). Therefore, genes found as multiple copies may outline their specific adaptations. Using BLAST (*17*), we identified six gene families with more than three paralogs. Comparing their copy numbers with those in other bacterial genomes, we found that five gene families (Tlc, SpoT, ProP, Sca, and AmpG) were

Table 1. Comparison between *R. conorii* and *R. prowazekii*.

Species	*R. conorii*	*R. prowazekii*
Genome size (bp)	1,268,755	1,111,523
G+C content (%)	32.4	29.0
ORFs	1374	834
tRNAs	33	33
rRNAs	3	3
Other RNAs	2	2
Pseudogenes	2	12
Coding content	81%	76%

[1]Information Génétique & Structurale, CNRS–AVENTIS UMR 1889, 31 chemin Joseph Aiguier, 13402 Marseille Cedex 20, France. [2]Unité des Rickettsies, Université de la Méditerranée, Faculté de Médecine, CNRS UMR 6020, 27 boulevard Jean Moulin, 13385 Marseille Cedex 05, France. [3]Genoscope and CNRS UMR 8030, 2 rue Gaston Crémieux, 91000 Evry, France. [4]Institut Pasteur, Unité des Interactions Bactéries-Cellules, 28 rue du Docteur Roux, 75724 Paris, Cedex 15, France.

*To whom correspondence should be addressed. E-mail: Jean-Michel.Claverie@igs.cnrs-mrs.fr; Didier.Raoult@medecine.univ-mrs.fr

significantly overrepresented ($P < 0.05$, Fisher's exact test) in *Rickettsia* species [Web table 1 (*10*)]. Adenosine triphosphate (ATP)/adenosine diphosphate (ADP) translocases are known to be unique to *Rickettsia* spp. and *Chlamydia* spp. among bacteria and may be of plant origin (*18*). This gene allows the importation of ATP from the infected host cell. Five copies were found in *R. conorii* and *R. prowazekii*. Also, four SpoT copies were found in *R. conorii* and *R. prowazekii* (*19*). The SpoT protein hydrolyzes the nucleotide (p)ppGpp, also known as alarmone. This compound plays a major role in processes related to starvation in various bacteria (*20*). Alarmone also initiates the expression of virulence genes in *Legionella pneumophila* (*21*), the production of antibiotics in *Streptomyces coelicolor* (*22*), and the change in the cell density of *Myxococcus xanthus* (*23*). The four copies of SpoT may be related to the adaptation of *Rickettsia* to long starvation periods in pausing ticks or louse feces. Both *Rickettsia* species have significantly large numbers of ProP proline/betain transporter paralogs: 11 and 7 ORFs for *R. conorii* and *R. prowazekii*, respectively. In many organisms, including bacteria and plants, proline transporters play critical roles in the response to osmotic changes in the environment (*24*). *Leishmania donovani*, an intracellular protozoan parasitizing both arthropod and mammalian cells, use temperature-regulated proline transporters to adapt to different host temperature conditions (*25*). The numerous ProP paralogs in *Rickettsia* might be linked to their adaptation to osmotic stress or to the temperature-dependent regulation of their virulence known as "reactivation" (*26*, *27*). *R. conorii* harbors five genes (four in *R. prowazekii*) encoding for outer membrane proteins of the Sca family, including rOmpA, which accounts for antigenic differences between *Rickettsia* species. *R. conorii* exhibits four copies of AmpG (three in *R. prowazekii*) that are likely to contribute to its natural resistance to β-lactam antibiotics. Finally, genes for the ATP-binding protein of multidrug resistance ABC transporters are present as four copies in *R. conorii* (three in *R. prowazekii*), but these numbers are not significantly higher than in other bacteria.

The genome of *R. conorii* exhibits a much higher density of interspersed repetitive DNA than that of *R. prowazekii* (Fig. 2). In the *R. conorii* genome, we identified 10 families (656 elements) of repeated DNA (*28*, *29*). Those repeats vary in size between 19 and 172 base pairs (bp) and constitute 3.2% of the entire genome. Overall, the repeat fraction of the genome is G+C rich (40%) and is in part responsible for the higher G+C content of *R. conorii* as compared with *R. prowazekii*. The distribution of the repeated elements is essentially random throughout the genome (Fig. 1). The quasiperfect colinearity maintained between the two *Rickettsia* genomes contrasts with the view that the multiplication of interspersed repeats promotes genomic rearrangements (*30*, *31*).

The analysis of the *R. conorii* putative coding regions revealed numerous cases of consecutive ORFs matching consecutive segments of a single longer ORF in other species, including *R. prowazekii*. Such gene fragmentations (e.g., internal stop codons) are usually associated with "pseudogenes." However, a truncated form of the outer-membrane protein (rOmpA) is normally expressed in *R. felis* (*32*). In addition, most of the split ORFs retained the statistical properties (coding potential and codon bias) of normal coding regions and a good similarity with intact protein orthologs. This prompted us to anno-

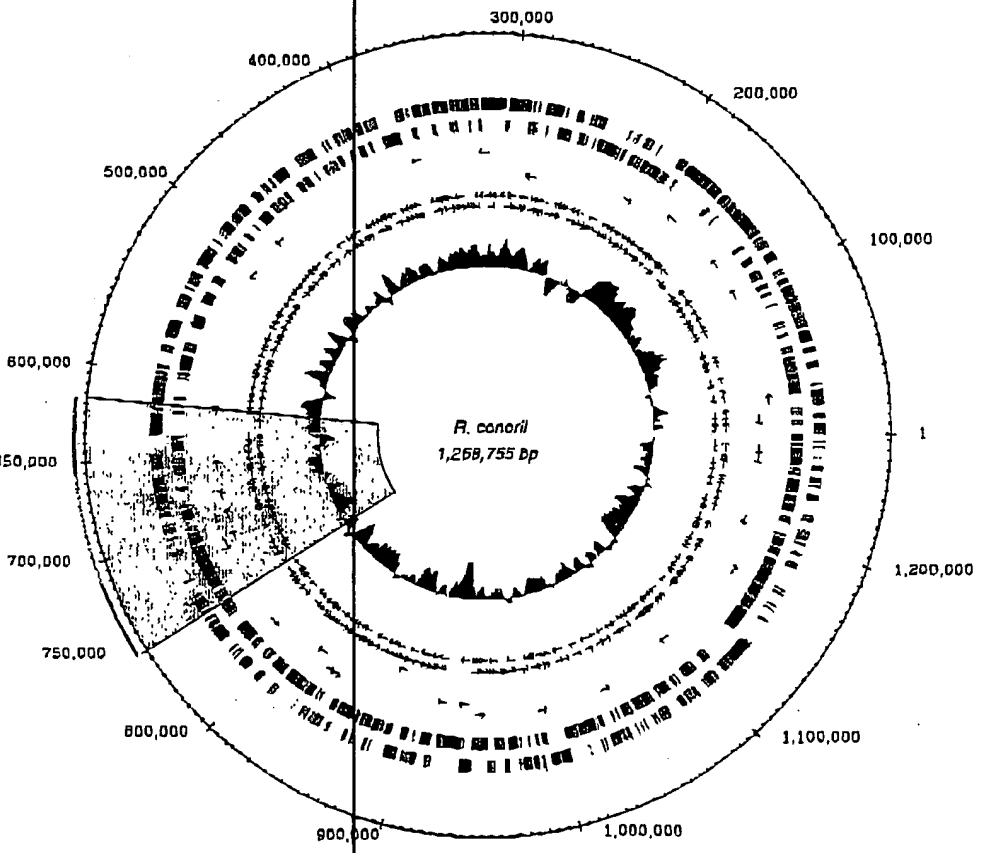


Fig. 1. Circular representation of the *R. conorii* genome (strain Malish 7). The outermost circle indicates the nucleotide positions. The second and third circles locate the ORFs on the plus and minus strands, respectively. Function categories are color-coded [see Web fig. 1 (*10*)]. The fourth and fifth circles locate tRNAs. The locations of three rRNAs are indicated by black arrows. The sixth and seventh circles indicate the locations of repeats. The eighth circle shows the G-C skew (G-C/G+C) with a window size of 10 kb. The region locally breaking the genome colinearity with *R. prowazekii* is indicated by a shaded sector. The four major genomic segments involved in this rearrangement are colored in blue, yellow, green, and red [see Fig. 3 and Web fig. 1 (*10*) for details].

tate these altered genes by the more neutral designation of "split genes," pending further experimental evidence. Thirty-seven split genes (*11*) (resulting in 105 total ORFs) were identified in *R. conorii* (Table 2). Among them, 14 have intact orthologs in *R. prowazekii*, 4 exhibit intact paralogs in *R. conorii*, and 19 have intact homologs in other prokaryotes. In *R. prowazekii*, we identified 11 split genes (resulting into 23 ORFs) that all have intact orthologs in *R. conorii*. By reverse transcriptase–polymerase chain reaction (RT-PCR) (*33*), we examined the detailed transcription pattern of all 37 *R. conorii* split genes [Web table 2 (*10*)]. We observed at least one transcript for 30 of 37 genes and RT-PCR products (all of the expected size) for 59 of 105 ORFs. All ORFs were transcribed for 11 of 31 genes, and the sole 5′ ORF was transcribed for 8 genes. These cases are consistent with the continued usage of the promoter of the original gene. However, the RT-PCR results on the other split genes suggested more complex transcription patterns. In seven cases, only the 3′ ORFs were found to be transcribed; in four cases, the 5′ and 3′ segments were found to be transcribed but not the middle segments; and in one case, only the middle segment was detected. Transcripts were much more likely to be detected for larger ORFs (≥70 residues) than for smaller ORFs (<70 residues) (55 of 78 versus 4 of 27; $P < 10^{-6}$, Fisher's exact test). These results suggest that these split genes might have retained some of their original

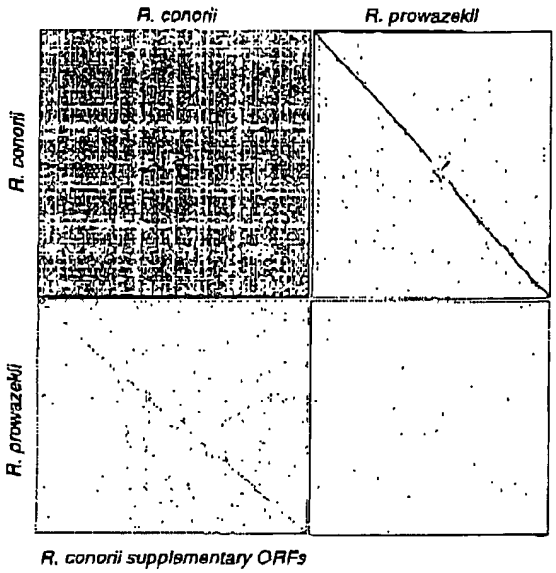


Fig. 2. Repeat density and colinearity of the *R. conorii* and *R. prowazekii* genomes. The two self-genome comparisons and the cross-genome comparison are presented in the upper left, upper right, and lower right panels. Each dot represents a high-scoring segment pair (HSP) identified by BLASTN (with a fixed database size parameter of 1 Mbp). Self matches are not shown. Red and black dots correspond to HSPs of *E* value < 10^{-4} and *E* value < 10^{-10}, respectively. In addition, the lower left panel presents the similarities of *R. conorii* 552 supplementary ORFs with the intergenic regions of *R. prowazekii*. Each dot represents an HSP detected by BLASTN (*E* value < 0.1). The black/red dots correspond to matches on the same/reverse strand, respectively. For the sake of readability, dot sizes are standard in all the panels and do not correspond to the actual size of the HSP.

Table 2. Numbers of ORFs (each split case was counted once). Numbers in parentheses indicate the corresponding number of intact ORFs. Rp, *R. prowazekii*; Rc, *R. conorii*.

Function	Orthologous ORFs			*R. conorii* ORFs in excess					*R. prowazekii* ORFs in excess		
	Reciprocal best match	Split in Rp	Split in Rc	Remnant in Rp	Remnant in Rp, split in Rc	Similarity in Rp (paralog)	Similarity in Rp (paralog), split in Rc	No remnant in Rp	No remnant in Rp, split in Rc	Remnant in Rc	Similarity in Rc (paralog)	No remnant in Rc
Replication	45			4	5 (2)	6		1				
Transcription	22											
Translation	92	2 (1)		1		2						
Protein modification	14						2 (1)					
Degradation of protein	16	2 (1)	4 (1)			3		2				
Transporters	52		3 (1)	7	6 (3)	1	5 (2)					
Regulatory functions	10		3 (1)	2		3	2 (1)		2 (1)			
Energy metabolism	69		3 (1)				2 (1)					1
Fatty acid and phospholipid	21		11 (3)	3			2 (1)			1		
Nucleotide metabolism	16				2 (1)							
Amino acid metabolism	6			2								
Cell envelope	61	5 (2)	4 (2)	1	2 (1)	2	2 (1)	1				
Metabolism of cofactors	27	2 (1)		2	4 (1)							
Cellular process	36			2								
Drug sensitivity	14	2 (1)	7 (2)	4	8 (2)	2		1				
Other category	14											
Unclassified	36			3	2 (1)	4		1			1	
Unknown	216	10 (5)	9 (3)	169		134	7 (2)	128	8 (2)	5	5	17
R. conorii total 1374 (1306)	767	11 (11)	44 (14)	200	29 (11)	157	22 (9)	134	10 (3)	–	–	–
R. prowazekii total 834 (822)	767	23 (11)	14 (14)	–	–	–	–	–	–	6	6	18

functions [as already discussed for SpoT genes in *R. prowazekii* (*19*)]. The complete assessment of the physiological significances of these genes will require a detailed characterization of their translation products in *Rickettsia*.

It has been previously suggested that most of the intergenic sequences in *R. prowazekii* consist of decayed genes that are no longer active but are not yet totally eliminated from the genome (*3*, *8*, *19*). Through a systematic survey, we identified noncoding remnant sequences for 229 ORFs (out of the 552 *R. conorii* supplementary genes) at their homologous locations in the *R. prowazekii* genome (Fig. 2). For example, the *R. conorii* gene (RC1273) for the outer membrane protein rOmpA is 6063 bp long and is located between the cell division protein FtsK (RC1274) and a hypothetical ORF (RC1272) (Fig. 3). *R. prowazekii* exhibits the orthologs for FtsK and the hypothetical ORF but not for rOmpA. Part of the intergenic sequence between the *R. prowazekii* orthologs exhibits a significant similarity to the rOmpA gene of *R. conorii*, although the remnant sequence identified in *R. prowazekii* (369 bp) contains several in-frame stop codons. Our comparative genome analysis thus strongly supports a model of massive gene decay in *R. prowazekii*.

We found one clear case of a gene decaying after its horizontal transfer from (or, less likely, to) *Chlamydia*. *R. conorii* exhibits a split form (RC0035-38) of the gene for a bifunctional folate synthesis protein described in Chlamydiae as composed of two distinct domains: 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase domain (HPPK) and dihydropteroate synthase domain (DHPS). Homologs of this enzyme with the same domain organization are found only in Chlamydiae, plants, and fungi. *R. prowazekii* exhibits remnant sequences corresponding to the *R. conorii* ORFs. The proximity of the *R. conorii* ORFs and Chlamydia genes supported by a phylogenetic tree analysis and the exclusive presence of this gene in *R. conorii* and *Chlamydia* species among known bacteria suggest a gene exchange between *Rickettsia* and *Chlamydia*, as proposed for ATP/ADP translocases (*18*). The alteration of this folate synthesis gene in both *Rickettsia* species—detected as a split gene in *R. conorii* and as a remnant sequence in *R. prowazekii*—suggests that significant changes of evolutionary constraints occurred after exchanging the gene with *Chlamydia*.

To investigate gene degradation in a wider set of *Rickettsia* species for which no sequence was available, we performed PCR assays on the genomic DNA of eight different *Rickettsia* species (*R. typhi*, *R. canadensis*, *R. helvetica*, *R. felis*, *R. australis*, *R. akari*, *R. rickettsii*, and *R. massiliae*) using primers derived from seven *R. conorii* and seven *R. prowazekii* supplementary genes [Web table 3 (*10*)]. With the primers corresponding to the seven *R. conorii* genes, two or more genomic segments were amplified in seven species [*R. rickettsii* (7 of 7), *R. australis* (6 of 7), *R. felis* (6 of 7), *R. helvetica* (5 of 7) *R. massiliae* (5 of 7), *R. akari* (3 of 7), and *R.*

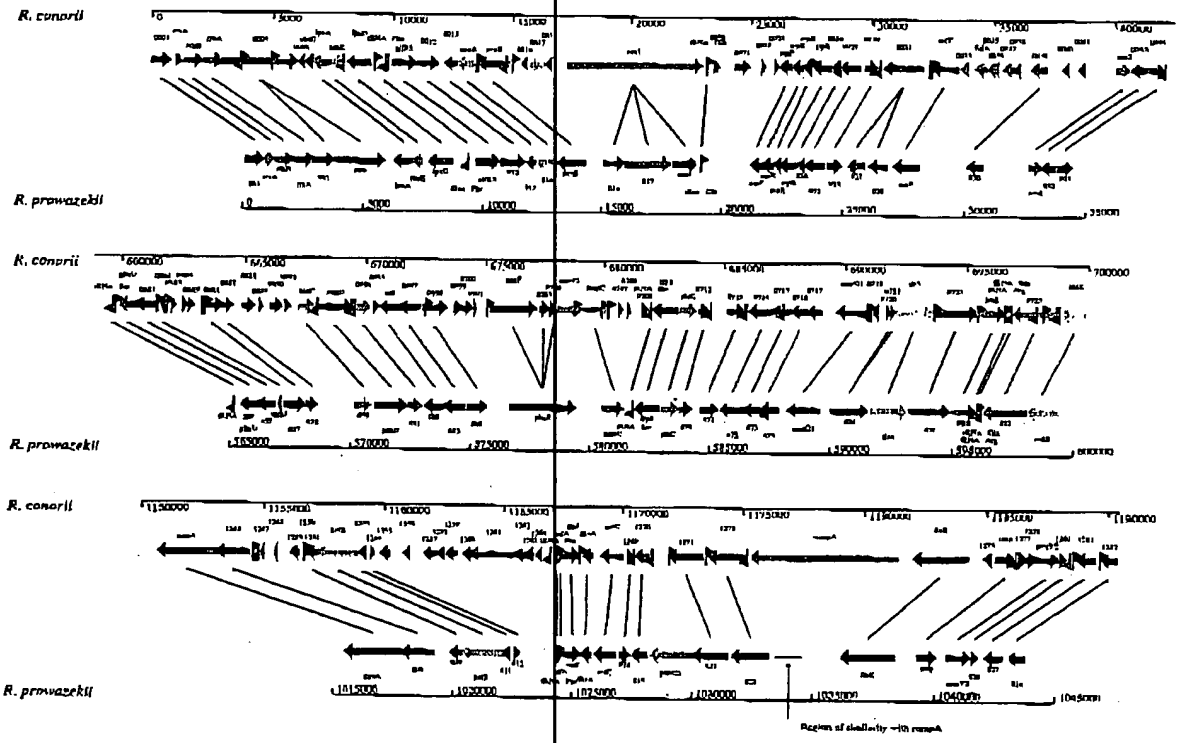


Fig. 3. Illustration of the colinearity. Three distinct segments from the *R. conorii* genome aligned with the homologous segments from the *R. prowazekii* genome are shown. These segments were chosen to show three types of gene alteration: split genes in *R. prowazekii* (top), a split gene in *R. conorii* (middle), and a gene remnant in *R. prowazekii* (bottom). A complete comparative map is available at (*10*).

canadensis (2 of 7)]. With the primers derived from the *R. prowazekii* genes, one or more genomic segments were amplified in four species [*R. typhi* (3 of 7), *R. australis* (1 of 7), *R. felis* (1 of 7), and *R. akari* (1 of 7)]. Thus, the supplementary genes observed in *R. conorii* and *R. prowazekii* do not originate from recent, species-specific, horizontal acquisitions, although the detailed pattern of PCR amplification does not exactly fit the standard classification of the *Rickettsia* genus [Web fig. 2 (*10*)]. Three out of the 14 supplementary genes were found in a split form (insertions/deletions generating stop codons) in one or more of the tested species. Thus, gene degradation appears to be a common feature of *Rickettsia*, targeting overlapping subsets of potentially dispensable genes while adapting to the selective pressures of different niches.

The few inversions/translocations locally breaking the otherwise perfect colinearity of the *R. conorii* and *R. prowazekii* genomes occur in the termination region of DNA replication. We identified several rearranged DNA segments, including a 38-kb segment containing 45 ORFs in *R. conorii* (Fig. 1). To date this inversion/translocation event within the phylogeny of *Rickettsia*, we used PCR on a set of primers designed from highly conserved adjacent sequences in the above eight species. A *R. prowazekii*–like arrangement was observed for *R. typhi*, whereas that of *R. conorii* was observed for *R. felis*, *R. rickettsii*, and *R. massiliae*. The result is consistent with the biphyletic division of *Rickettsia* and suggests that the genome rearrangement event would have happened relatively shortly after the initial divergence of the genus *Rickettsia*.

Genome reduction is thought to be a main force behind the evolution of parasitic and/or intracellular bacteria (*6*–*9*). The sequence of the *R. conorii* genome is consistent with this view, and *R. prowazekii* essentially appears as a subset of *R. conorii*. However, the genomes of *R. conorii* and *R. prowazekii* exhibit large differences in size, as well as in gene and G+C content, thus suggesting an adaptation to their specific niches rather than a simple model of random gene loss. Our analysis pointed out 137 *R. conorii* genes without any sequence similarity within the *R. prowazekii* genome. This provides an upper limit on the number of potential genes laterally acquired since their divergence, 40 to 80 million years ago (*34*). A single gene has its best match in eukaryotes (RC0781 to the NH_2-terminal part of yeast biotin-protein ligase), suggesting that *Rickettsia* have no particular tendency to evolve by acquiring genes from their hosts. Given their genetic isolation, it is tempting to postulate that *Rickettsia* had to rely on internal mechanisms such as duplication to acquire or modify some of the gene functions required for a better adaptation to their niche. We saw evidence of the very gradual nature of the genome reduction process by identifying all possible intermediates from intact ORFs: transcribed split ORFs, further split ORFs no longer transcribed, fully decayed but still recognizable ORFs, and complete gene disappearance. Similar mechanisms probably occur in the evolution of all bacterial species but have remained undetected because of more active recombination and a faster evolutionary rate (*35*).

References and Notes

1. D. Raoult, V. Roux, *Clin. Microbiol. Rev.* 10, 694 (1997).
2. P. Yagupsky, B. Wolach, *Clin. Infect. Dis.* 17, 850 (1993).
3. S. G. E. Andersson *et al.*, *Nature* 396, 133 (1998).
4. *R. conorii* (strain Malish 7) was grown and purified as previously reported (*36*). A preliminary library A (3-kb inserts cloned in pcDNAZ.1 with Bst XI adaptators) was constructed and one genomic equivalent was sequenced. After the pilot study, we generated two other libraries of different insert sizes, F (5 kb) and G (10 kb) cloned in the same vector as A. Plasmid clones were sequenced at both ends of the insert with flanking vector sequences as primers. Dye primer reactions were analyzed on a LI-COR 4200 L (library A) and dye-terminator reactions were analyzed on a capillary ABI3700 (libraries F and G). The whole genome sequence assembly was performed by means of the PHRED and PHRAP software packages (*37*, *38*). A total of 9603 (6.4X), 5855 (2X), and 2352 (0.8X) end sequences (X's indicate genome equivalents), respectively, from libraries A, F, and G were incorporated into contigs. Cloning gaps and several sensitive regions were resolved or confirmed by sequencing duplicated PCR products. The final sequence includes 99.98% positions with PHRED scores over 40. None of the remaining 278 lower quality positions were found to coincide with a stop codon in our final annotation of ORFs nor were they related to split genes. Nucleotides in the *R. conorii* genome were numbered according to the *R. prowazekii* genome, where position one corresponds to the predicted origin of replication. The coherence of the assembly was verified by comparison to the Dsg I and Sma I restriction site pattern previously obtained by pulse-field gel electrophoresis (*36*).
5. ORFs were delineated with a bioinformatics protocol similar to the procedure used to annotate the *R. prowazekii* genome sequence (*3*). Following Andersson *et al.*, we classified candidate ORFs in three categories: (i) ORFs at least 150 nucleotides (nt) long associated with high coding potential, (ii) ORFs at least 150 nt long exhibiting significant sequence similarity [*E* value < 0.005, with BLASTP (*39*) against the National Center for Biotechnology Information nonredundant protein database], and (iii) the remaining ORFs that are at least 300 nt long with neither coding potential nor similarity. Coding potential was evaluated with the SelfID program (*40*). Two ORFs for ribosomal proteins shorter than 50 amino acid residues were identified later. After this first pass, pairs of ORFs overlapping by more than 50% of the size of the shorter ORF were further handled as follows: The ORF exhibiting the highest scoring database matches was kept, irrespective of length. In case of no matches occurring, the longest ORF was selected. The final annotation contains 1374 ORFs. Genes for transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and other RNAs were identified by BLASTN similarity searches against multiple databases. Two interrupted ORFs in *R. conorii* were annotated as pseudogenes on the basis of their colinearity and sequence similarity with two pseudogenes already annotated as such in *R. prowazekii*. Tentative ORF functions were assigned on the basis of sequence similarity against protein sequence databases [KEGG (*41*), SWISSPROT (*42*), and the nonredundant protein database] and domain/motif databases [Pfam (*43*) and PROSITE (*44*)]. Orthologous versus paralogous relationships were identified with multiple sequence alignments and neighbor-joining trees constructed by ClustalW (*45*). The functional classification was based on the scheme provided by the *R. prowazekii* genome project (*3*) as well as that provided by KEGG.
6. S. G. E. Andersson, C. G. Kurland, *Trends Microbiol.* 6, 263 (1998).
7. J. O. Andersson, S. G. E. Andersson, *Curr. Opin. Genet. Dev.* 9, 664 (1999).
8. ———, *Mol. Biol. Evol.* 16, 1178 (1999).
9. S. T. Cole *et al.*, *Nature* 409, 1007 (2001).
10. Supplementary Web material is available on *Science* Online at www.sciencemag.org/cgi/content/full/293/5537/2093/DC1.
11. We introduce the concept of split genes to define the cases where a given protein sequence usually encoded as a single ORF appears divided into a set of two or more colinear ORFs (>150 nt) on the same strand of the chromosome of another species. Most (48 of 68 for *R. conorii* and 11 of 12 for *R. prowazekii*) of the newly introduced stop codons are followed by the putative initiation codon of the next ORF within a distance of 150 nt. Split genes were identified with a program originally developed to extract evolutionarily conserved gene strings (*46*).
12. Noncoding remnant sequences in genome *A* corresponding to bona fide genes in the other genome *B* were identified as follows (*A* and *B* denote *R. conorii* or *R. prowazekii*). For every given ORF index gene (g_i) only found in genome *B* (supplementary genes), the colinearity of adjacent genes was used to delineate their homologous noncoding regions g_i' in genome *A*. g_i' was then diagnosed as containing a remnant sequence of gene g_i when the segment most similar to g_i was found within g_i' after a FASTA (*47*) search against the entire genome *A*. Additional cases were identified by comparing g_i and g_i' with LALIGN (*48*) (with a similarity threshold of *E* value $<10^{-10}$) and a subsequent analysis by dot-plot.
13. S. Kalman *et al.*, *Nature Genet.* 21, 385 (1999).
14. V. Roux, *Rickettsiae and Rickettsial Diseases at the Turn of the Third Millennium*, D. Raoult, P. Brouqui, Eds. (Elsevier, Paris, France, 1999), p. 52.
15. N. Teysseire, C. Chiche-Portiche, D. Raoult, *Res. Microbiol.* 143, 821 (1992).
16. P. Cossart, *Curr. Opin. Cell Biol.* 7, 94 (1995).
17. Gene families were initially identified with the single linkage clustering method with BLASTP (with a threshold of *E* value $< 10^{-30}$). Out of the resulting clusters, six were composed of four or more ORFs: ATP/ADP translocases (five ORFs), proline/betaine transporters ProP (seven ORFs), (p)ppGpp 3-pyrophosphohydrolase SpoT (four ORFs), AmpG proteins (four ORFs), outer membrane proteins (four ORFs), and multidrug-resistant ABC transporter ATP-binding proteins (including alkaline protease secretion proteins) (five ORFs). According to the additional PSI-BLAST searches and careful examinations of the results, we identified four more ORFs for ProP, one more ORF for SpoT, and one more ORF for the Sca proteins.
18. Y. I. Wolf, L. Aravind, E. V. Koonin, *Trends Genet.* 15, 173 (1999).
19. J. O. Andersson, S. G. E. Andersson, *Mol. Biol. Evol.* 18, 829 (2001).
20. G. Schreiber, E. Z. Ron, G. Glaser, *Curr. Microbiol.* 30, 27 (1995).
21. B. K. Hammer, M. S. Swanson, *Mol. Microbiol.* 33, 721 (1999).
22. R. Chakraburtty, M. Bibb, *J. Bacteriol.* 179, 5854 (1997).
23. B. Z. Harris, D. Kaiser, M. Singer, *Genes Dev.* 12, 1022 (1998).
24. T. Leisinger, in Escherichia coli *and* Salmonella, F. C. Neidhardt *et al.*, Eds. (American Society for Microbiology, Washington, DC, 1996), pp. 434–440.
25. S. Mazareb, Z. Y. Fu, D. Zilberstein, *Exp. Parasitol.* 91, 341 (1999).

26. P. F. Policastro, U. G. Munderloh, E. R. Fischer, T. E. D. Hackstadt, *J. Med. Microbiol.* 46, 839 (1997).
27. S. F. Hayes, W. Burgdorfer, *Infect. Immun.* 37, 779 (1982).
28. The repeated elements in the *R. conorii* genome were initially identified on the basis of the direct self-comparison of the genomic DNA by the BLASTN program (*E* value < 10^{-4}). The BLASTN result was then analyzed by the repeat identification program Mreca (49). An exhaustive analysis (types, locations, insertion in ORFs, etc.) of the repeats identified in the *R. conorii* genomic sequence is available under the *Rickettsia* database section at http://igs-server.cnrs-mrs.fr/RicBase/.
29. Ten repeat families include the previously identified *Rickettsia* palindromic element described in H. Ogata *et al.*, *Science* 290, 347 (2000), and H. Ogata, S. Audic, J.-M. Claverie, *Science* 291, 252 (2001).
30. D. R. Leach, *Bioessays* 16, 893 (1994).
31. P. M. Sharp, D. R. Leach, *Mol. Microbiol.* 22, 1055 (1996).
32. D. Raoult *et al.*, *Emerg. Infect. Dis.* 7, 73 (2001).
33. The presence of transcripts associated with all *R. conorii* split genes was assessed by RT-PCR with primers designed from their predicted coding regions. In addition, we tested five genes (RC0873, RC0426, RC0888, RC0374, and RC0909) of *R. conorii* with no functional counterparts in *R. prowazekii*. Primer sequences and expected amplicon sizes are listed in Web table 2 (10). RNA was prepared as follows: A suspension of fresh *R. conorii* was adjusted to 10^8/ml and the bacteria were separated from cells with a sucrose gradient. RNA was extracted with the RN easy Midi kit (Qiagen, Hilden, Germany) as recommended by the manufacturer. RT-PCR was performed on the resulting RNA with the OneStep RT-PCR kit (Qiagen) according to the manufacturer's protocol. Reverse transcription and amplifications were performed in PTC-200 thermocyclers (MJ Research, Watertown, MA) with 40 PCR cycles. RT-PCR products were run in 1% agarose gels, stained with ethidium bromide, and revealed on an ultraviolet box. All amplicons exhibited the expected size. Once RT-PCR was performed, all primer pairs were validated by positive PCR reactions on genomic DNA. The absence of genomic DNA contamination in the RNA preparation was controlled by attempting PCR (with the Gibco-Brl Elongase polymerase) with the primers for ORF RC0909 (WASP-like gene) and ORF RC0843 (comF gene). The PCR assay was negative in both cases.
34. The divergence time was estimated on the basis of 16S rRNA sequences, assuming a relatively constant rate of evolution of 1 to 2% per 50 million years (50).
35. E. J. Fell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 98, 182 (2001).
36. V. Roux, D. Raoult, *J. Bacteriol.* 175, 4895 (1993).
37. B. Ewing, L. Hillier, M. C. Wendl, P. Green, *Genome Res.* 8, 175 (1998).
38. B. Ewing, P. Green, *Genome Res.* 8, 186 (1998).
39. S. F. Altschul *et al.*, *Nucleic Acids Res.* 25, 3389 (1997).
40. S. Audic, J.-M. Claverie, *Proc. Natl. Acad. Sci. U.S.A* 95, 10026 (1998).
41. M. Kanehisa, S. Goto, *Nucleic Acids Res.* 28, 27 (2000).
42. A. Bairoch, R. Apweiler, *Nucleic Acids Res.* 28, 45 (2000).
43. A. Bateman *et al.*, *Nucleic Acids Res.* 28, 263 (2000).
44. K. Hofmann, P. Bucher, L. Falquet, A. Bairoch, *Nucleic Acids Res.* 27, 215 (1999).
45. J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acids Res.* 22, 4673 (1994).
46. H. Ogata, W. Fujibuchi, S. Goto, M. Kanehisa, *Nucleic Acids Res.* 28, 4021 (2000).
47. W. R. Pearson, D. J. Lipman, *Proc. Natl. Acad. Sci. U.S.A.* 85, 2444 (1988).
48. X. Huang, W. Miller, *Adv. Appl. Math.* 12, 373 (1991).
49. C. Notredame, *Bioinformatics* 17, 373 (2001).
50. H. Ochman, S. Elwyn, N. A. Moran, *Proc. Natl. Acad. Sci. U.S.A.* 96, 12638 (1999).

51. We thank P. Sansonetti for strongly supporting the initiation of this project and F. Gosse for reading the manuscript. Supported in part by the Hoechst, Marion, Roussel research fund (grant no. FRHMR1/9777) and by Aventis Pharma, France. A more comprehensive database on the *R. conorii* genome and proteome is available at http://igs-server.cnrs-mrs.fr/RicBase/.

9 April 2001; accepted 23 July 2001

Role of Nonimmune IgG Bound to PfEMP1 in Placental Malaria

Kirsten Flick,[1] Carin Scholander,[1] Qijun Chen,[1] Victor Fernandez,[1] Bruno Pouvelle,[2] Jurg Gysin,[2] Mats Wahlgren[1]*

Infections with *Plasmodium falciparum* during pregnancy lead to the accumulation of parasitized red blood cells (infected erythrocytes, IEs) in the placenta. IEs of *P. falciparum* isolates that infect the human placenta were found to bind immunoglobulin G (IgG). A strain of *P. falciparum* cloned for IgG binding adhered massively to placental syncytiotrophoblasts in a pattern similar to that of natural infections. Adherence was inhibited by IgG-binding proteins, but not by glycosaminoglycans or enzymatic digestion of chondroitin sulfate A or hyaluronic acid. Normal, nonimmune IgG that is bound to a duffy binding–like domain β of the *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) might at the IE surface act as a bridge to neonatal Fc receptors of the placenta.

Malaria infection with *P. falciparum* during pregnancy is an important cause of maternal morbidity and mortality. It may induce premature delivery, spontaneous abortion, or lead to a low birth weight (*1*, *2*). Infections often cause more severe symptoms in primiparous than in multiparous women. The incidence of placental malaria similarly diminishes with increasing parity (*3*, *4*), probably due to the acquisition of immunity to the infecting parasites (*5*, *6*).

IEs are not passed from the mother to the fetus, but accumulate in the placenta which can experience high parasite densities (>50% IEs) while the peripheral circulation is almost free of IEs. Placental malaria may thus be caused by IEs that are selected for and expanded on receptors only present in the placenta (*7*–*10*), as opposed to those in other vascular beds.

Certain strains of *P. falciparum* bind nonimmune immunoglobulins onto the surface of the host erythrocyte, a fact that made us investigate their role in sequestration, in particular the possibility that IgG could bridge the IEs to Fc-receptors present in the placenta. We thus examined the frequency of IgG- (and IgM-) binding IEs accumulated during pregnancy in the placenta. Small pieces of snap-frozen placental tissues were obtained from six malaria-infected Cameroonese women after approved consent. The parasitemia of the placentas ranged from three to 23% (Fig. 1A), and all of them were classified as having active or active-chronic infections (*11*). IgG-binding IEs (Fig. 1B) were found in all of the placentas (10 to 75% IgG positive, mean 44%), whereas IgM-binding IEs (Fig. 1C) were more rare (2 to 34%, mean 18%) (*12*). IEs attached to the syncytiotrophoblasts bound only IgG (20 to 80%, mean 50%), except those of placenta CP42DJ, where the number of IgG-binding IEs was equal to that of the IgM-binding IEs. By studying the Ig-binding phenotype of IEs eluted from the placentas (*13*, *14*), we confirmed that a majority of parasites causing active placental infection bound IgG (Table

Table 1. The phenotype of *P. falciparum* infecting the human placenta. IEs were eluted from the infected placentas and scored for their immunoglobulin binding.

Placenta	IgG-binding IE/IE tested*	Percent
CP24	18/23	78
CP42	14/14	100
CP42DJ†	–	–
CP193	205/276	74
CP939	32/247	13
CP940†	–	–

*Fractions of the IEs were also studied for their capacity for binding to Sca1D cells including inhibition with soluble CSA (100 μg/ml) and treatment with CSAnase ABC (0.5 U). About 50% of the IEs were specific for CSA (*11*). The populations of parasites studied for CSA-binding were not identical to those studied for immunoglobulin-binding since not all eluted IEs were scored in the CSA-assays. †The number of eluted parasites obtained from the placentas CP42DJ and CP940 were not sufficient for the assay.

[1]Microbiology and Tumor Biology Center (MTC), Karolinska Institutet and Swedish Institute for Infectious Disease Control, Box 280, S-171 77 Stockholm, Sweden. [2]Unité de Parasitologie Expérimentale, Faculté de Médecine, Université de la Méditerranée (Aix-Marseille II), 13385 Marseille Cedex 5, France.

*To whom correspondence should be addressed. E-mail: Mats.wahlgren@smi.ki.se

Biochem. J. (2002) 363, 313–319 (Printed in Great Britain)

Folate synthesis in higher-plant mitochondria: coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities

Jean-Marie MOUILLON, Stéphane RAVANEL, Roland DOUCE and Fabrice RÉBEILLÉ[1]

Laboratoire de Physiologie Cellulaire Végétale, UMR 5019, CEA/CNRS/Université Joseph Fourier, Département de Biologie Moléculaire et Structurale, CEA-Grenoble, 17 rue des Martyrs, F-38054 Grenoble Cedex 9, France

The plant enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase (HPPK/DHPS) is a mitochondrial bifunctional protein involved in tetrahydrofolate synthesis. The first domain (HPPK) catalyses the pyrophosphorylation of 6-hydroxymethyl-7,8-dihydropterin (dihydropterin) by ATP, leading to 6-hydroxymethyl-7,8-dihydropterin pyrophosphate (dihydropterinPP$_i$) and AMP. The second domain (DHPS) catalyses the next step, i.e. the condensation of *p*-aminobenzoic acid (*p*-ABA) with dihydropterinPP$_i$ to give 7,8-dihydropteroate (dihydropteroate) and PP$_i$. In the present article we studied the coupling between these two reactions. Kinetic data obtained for the HPPK domain are consistent with an ordered Bi Bi mechanism where ATP binds first and dihydropterinPP$_i$ is released last, as proposed previously for the monofunctional *Escherichia coli* enzyme. In the absence of *p*-ABA, AMP and dihydropterinPP$_i$ accumulate and negatively regulate the reaction. In the presence of *p*-ABA, the rates of AMP and dihydropteroate synthesis are similar, indicating a good coupling between the two reactions. DihydropterinPP$_i$, an intermediate of the two reactions, never accumulates in this situation. The high specific activity of DHPS relative to HPPK, rather than a preferential channelling of dihydropterinPP$_i$ between the two catalytic sites, could explain these kinetic data. The maximal velocity of the DHPS domain is limited by the availability of dihydropterinPP$_i$. It is strongly feedback-inhibited by dihydropteroate and also dihydrofolate and tetrahydrofolate monoglutamate, two intermediates synthesized downstream in the folate biosynthetic pathway. Thus the HPPK domain of this bifunctional protein is the limiting factor of the overall reaction, but the DHPS domain is a potential key regulatory point of the whole folate biosynthetic pathway.

Key word: C1 metabolism.

INTRODUCTION

Plants and micro-organisms, in contrast to animals, are able to synthesize tetrahydrofolate *de novo*. From 6-hydroxymethyl-7,8-dihydropterin (hereon referred to as dihydropterin), this biosynthetic pathway requires the sequential operation of five enzymes (for a review, see [1]). The enzymes catalysing the first three steps are absent in animals. These enzymes are therefore potential targets for antimicrobial or herbicide drugs. The 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK; reaction 1) and the 7,8-dihydropteroate synthase (DHPS; reaction 2) catalyse the first two steps of this pathway, leading to the formation of dihydropteroate.

$$\text{ATP} + \text{dihydropterin} \rightarrow \text{AMP} + \text{dihydropterinPP}_i \quad (1)$$

$$\text{DihydropterinPP}_i + p\text{-ABA} \rightarrow \text{dihydropteroate} + \text{PP}_i \quad (2)$$

where dihydropterinPP$_i$ is 6-hydroxymethyl-7,8-dihydropterin pyrophosphate and *p*-ABA is *p*-aminobenzoic acid. The DHPS activity (reaction 2) is the target of sulphonamides, a family of molecules that are potent antibiotics. These chemicals are *p*-ABA analogues that are recognized by the enzyme as alternative substrates [2,3] and their effects in bacteria and protozoa have been studied intensively [4–8]. DHPS from prokaryotes is a dimeric protein with identical subunits of about 30–35 kDa [9,10]. The three-dimensional structures of the *Escherichia coli* and *Staphylococcus aureus* DHPS proteins have been solved, giving new insights into how substrates and inhibitors bind and react in the catalytic site of the protein [11,12]. The protein belongs to the triose phosphate isomerase (TIM) barrel group and kinetic and crystallographic studies suggest a random binding of the two substrates and a half-site reactivity (the substrates bind to only one monomer) [12].

On the other hand, there are few data concerning the kinetic behaviour of the HPPK-catalysed reaction. This reaction was generally measured in association with DHPS activity [13,14], and the regulation of this pyrophosphoryl-transferring enzyme remains to be determined. In prokaryotes, the HPPK protein is a monomer of 18 kDa [13]. The three-dimensional structures of the *E. coli* and *Haemophylus influenzae* enzymes have been determined [15–17], showing that the monomer consists of a single globular α/β domain. The positions of the substrates in the catalytic site are compatible with a mechanism involving a nucleophilic attack from the hydroxyl group of the pterin molecule to the β-phosphate of ATP, followed by a direct transfer of the PP$_i$ group from ATP to the 6-hydroxymethyl side chain of dihydropterin. Two Mg^{2+} ions, co-ordinated with the carboxyl groups of two conserved aspartate residues, are involved in this transfer and stabilize the transition state.

In contrast with the situation found in prokaryotes, HPPK and DHPS activities are associated in all the eukaryotes studied so far. Indeed, in the protozoa *Plasmodium falciparum* and *Toxoplasma gondii* a bifunctional protein [18,19] supports both activities. The situation is even more complex in the sporozoan *Pneumocystis carinii*, where the enzyme is a trifunctional polypeptide supporting dihydroneopterin aldolase, HPPK and DHPS activities [20]. Dihydroneopterin aldolase catalyses the conversion of dihydroneopterin into dihydropterin, the substrate of HPPK activity. In plants, the DHPS and HPPK activities are

Abbreviations used: *p*-ABA, *p*-aminobenzoic acid; HPPK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase; DHPS, 7,8-dihydropteroate synthase; dihydropterin, 6-hydroxymethyl-7,8-dihydropterin; dihydropterinPP$_i$, 6-hydroxymethyl-7,8-dihydropterin pyrophosphate.

[1] To whom correspondence should be addressed (e-mail fabrice.rebeille@cea.fr).

also supported by a bifunctional enzyme located in the matrix space of mitochondria [21,22]. In fact, all the five steps required for tetrahydrofolate polyglutamate synthesis in plants were localized in mitochondria [21], which strongly suggests that these organelles are a major site for folate synthesis in these organisms.

In previous studies [22] we have purified and determined the primary sequence of the bifunctional HPPK/DHPS from pea leaves. The molecular data (the presence of a mitochondrial transit peptide, a single-copy gene) confirmed the biochemical evidence for this protein being localized in the mitochondria. Preliminary kinetic studies of the DHPS activity suggested a random bireactant system strongly retro-inhibited by dihydropteroate, a product of the reaction (K_i, 8–10 μM [22]). However, the catalytic properties of HPPK and its coupling with DHPS have not been studied yet. The biochemical characterization of the native HPPK/DHPS from higher-plant mitochondria is hampered by the difficulty in obtaining sufficient amounts of protein. Indeed, it represents only 0.04–0.06% of the soluble proteins of the mitochondrial matrix [22]. An alternative approach is to produce a recombinant enzyme in a prokaryotic host, such as *E. coli*. In this report, we describe for the first time the catalytic properties and coupling of the HPPK and DHPS activities of the plant recombinant bifunctional enzyme.

MATERIALS AND METHODS

Materials

Pteroate and pterin were obtained from Sigma. PterinPP$_i$ was obtained from Schircks Laboratory (Jona, Switzerland). These products were reduced as dihydro- compounds as described by Scrimgeour [23]. They were quantified by their typical absorption spectra [24,25] and MS analysis, then stored at −80 °C in flasks saturated with argon.

Expression of the recombinant HPPK/DHPS in *E. coli*

The cDNA fragment encoding the mature pea leaf HPPK/DHPS [22] was ligated into the appropriate sites of a modified version of the plasmid pET-3a (Stratagene), provided kindly by Professor D. Macherel (University of Angers, Angers, France). This vector contained the nucleotide sequence of *argU*, the gene encoding a tRNA^{Arg} that is only rarely found in bacteria. Indeed, the HPPK/DHPS protein contains 19 arginine residues, a situation that could strongly limit its prokaryotic expression. The resulting pET-HPPK/DHPS vector was used to transform the BL21 (DE3) *E. coli* strain (Stratagene). The bacteria were then grown at 37 °C in 2 litres of Luria–Bertani medium supplemented with 100 μg/ml carbenicillin. The expression of recombinant HPPK/DHPS protein was induced at 18 °C by the addition of 4 μM isopropyl β-D-thiogalactoside. Cells were harvested 48 h later by centrifugation. Most of the recombinant protein was produced as inclusion bodies. However, a small part remained soluble. The presence of the recombinant HPPK/DHPS was monitored by SDS/PAGE analysis [26] and by Western blotting using rabbit polyclonal antibodies raised against the recombinant enzyme purified from inclusion bodies (Elevage Scientifique des Dombes, Chatillon-Chalaronne, France).

Purification of the recombinant HPPK/DHPS

Cell pellets from 1 litre cultures were incubated at 37 °C for 15 min in 100 ml of a medium containing 2 mM EDTA, 10 μg/ml lysozyme and 50 mM Tris/HCl (pH 8.0). The cells were disrupted by three consecutive freeze/thaw cycles followed by a 10 min ultrasonic period at 4 °C. All subsequent steps were performed at 4 °C. The suspension was centrifuged for 15 min at 15000 *g* to pellet membrane fragments and inclusion bodies. Lipids and small vesicles were removed by ultra-centrifugation for 1 h at 30000 *g*. The resulting supernatant containing soluble recombinant HPPK/DHPS was then concentrated 10-fold by ultrafiltration through a 100 kDa-cutoff membrane (Amicon). This procedure removed the two bacterial HPPK (18 kDa) and DHPS (70 kDa) proteins present in vanishing amounts. The remaining protein extract was thereafter loaded on to a 0.5 cm × 5 cm folate–agarose affinity column (Sigma) that had been equilibrated previously with buffer I [0.1 M KH_2PO_4, pH 7.5/10% (v/v) glycerol]. The flow rate was 0.3 ml/min. The column was washed for 45 min with buffer I, then eluted with a linear gradient (0–100%) of buffer II [0.1 M KH_2PO_4, pH 8.5/10% (v/v) glycerol/1 mM folate]. Fractions containing the purified recombinant HPPK/DHPS were dialysed against buffer I and concentrated to a final volume of 800 μl. The N-terminal part of the protein was analysed as described previously [22]. The obtained sequence (MFHTAPNSSI) matched the one determined with the native enzyme [22], thus confirming its identity. The protein was stored at −80 °C in 10% (v/v) glycerol until use.

Determination of HPPK and DHPS activities

These activities were estimated at 30 °C. All the solutions were maintained under a stream of argon to minimize the oxidation of the various dihydropterin substrates. The HPPK + DHPS activity (total activity, i.e. reaction 1 + reaction 2) was estimated according to [21]. The standard reaction medium (medium A) contained, in a total volume of 120 μl: 40 mM Tris (pH 8.0), 20 mM 2-mercaptoethanol, 10 mM $MgCl_2$, 200 μM ATP and various amounts of the purified protein. [^{14}C]*p*-ABA (2 μl, 2 mM, 1.85 GBq·mmol^{-1}; ICN Biomedicals) was added to the assay medium, and then the reaction was started by the addition of 100 μM dihydropterin. After various incubation periods, 100 μl of the assay medium were injected into a reversed-phase HPLC system (Merck 655A-11 Liquid Chromatograph, equipped with a Shandon, Zorbax ODS Z225 5 μm column) coupled with a Berthold (LB 506D) scintillation counter, as described earlier [22]. The HPLC conditions were: solvent A, 0.1 M sodium acetate, pH 6; solvent B, acetonitrile; solvent B increased linearly by 0.8%/min at a flow rate of 1 ml·min^{-1}. Within these experimental conditions, excess [^{14}C]*p*-ABA was not retained in the column, whereas [^{14}C]dihydropteroate, the final product of the reaction, was eluted after 19 min of chromatography.

The DHPS activity (reaction 2) was measured in medium A (final volume, 120 μl) devoid of ATP [22]. For a standard assay, 2 μl of 2 mM [^{14}C]*p*-ABA (1.85 GBq·mmol^{-1}) were added in the medium, and the reaction was started by the addition of 100 μM dihydropterinPP$_i$. After various times of incubation the [^{14}C] dihydropteroate formed was estimated as described above.

The HPPK activity (reaction 1) was estimated through the formation of either [2,8-^{3}H]AMP in the presence of [2,8-^{3}H]ATP (Amersham Bioscience) or β-[^{32}P]dihydropterinPP$_i$ in the presence of [γ-^{32}P]ATP (Amersham Bioscience). The standard reaction medium contained, in a total volume of 120 μl: 40 mM Tris (pH 8.0), 20 mM 2-mercaptoethanol, 10 mM $MgCl_2$ and 200 μM [2,8-^{3}H]ATP (1.85 GBq·mmol^{-1}) or 200 μM [γ-^{32}P]ATP (1.85 GBq·mmol^{-1}) and various amounts of the purified recombinant enzyme. Then the reaction was started by the addition of 100 μM dihydropterin. After various periods of incubation, the products of the reaction were analysed by the HPLC system described above. The HPLC conditions were: solvent A, 50 mM KH_2PO_4, pH 6.0/5 mM tetrabutylammonium phosphate/4% (v/v) acetonitrile; solvent B, acetonitrile; solvent B increased linearly by 0.8%/min at a flow rate of 1 ml·min^{-1}. Within

these experimental conditions, [2,8-^{3}H]AMP and [2,8-^{3}H]ATP (or [γ-^{32}P]ATP) were eluted after 14 and 32 min of chromatography, respectively, whereas β-[^{32}P]dihydropterinPP$_i$ was eluted after 15 min. We verified in separate control experiments that non-enzymic formation of [2,8-^{3}H]AMP was negligible during the course of the experiment.

RESULTS

Expression and purification of the recombinant HPPK/DHPS

The recombinant HPPK/DHPS was expressed almost exclusively as inclusion bodies in *E. coli*. However, with the culture conditions described in the Materials and methods section (i.e. low temperature, low isopropyl β-D-thiogalactoside concentration) DHPS and HPPK activities were higher in the transformed cells than in the control, indicating that some of the recombinant protein was soluble and active. Although the bifunctional HPPK/DHPS was not clearly visible on SDS/PAGE analysis of soluble proteins (Figure 1A, lane 2), the antibodies raised against HPPK/DHPS recognized a band of approx. 53000 Da, corresponding to the molecular mass of the plant enzyme (Figure 1B, lane 1). This band was not present in non-transformed cells (result not shown). Unfortunately, all our attempts to improve the yield of recombinant enzyme production (expression as a fusion protein with thioredoxin, co-expression with GroES and GroEL, and expression in the yeast *Pichia pastoris*) were unsuccessful. Nevertheless, the production of the recombinant HPPK/DHPS in *E. coli* was a good alternative compared with the time-consuming purification of the enzyme from isolated plant mitochondria [22]. Indeed, the recombinant HPPK/DHPS enzyme was purified in a single step by affinity chromatography on a folate–agarose column, leading to a final enzyme preparation apparently devoid of contaminants (Figure 1A, lane 3). Sequencing of the N-terminal part of the protein (see the Materials and methods section) further confirmed the identity of the enzyme. According to our protocol, about 300 μg of purified HPPK/DHPS could be obtained from 1 litre of *E. coli* cell culture.

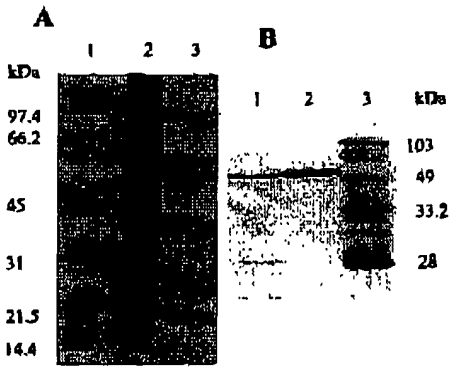


Figure 1 Purification and immunodetection of the recombinant HPPK/DHPS expressed in *E. coli*

(A) Coomassie Brilliant Blue-stained SDS/PAGE of the *E. coli* proteins during the course of purification. Lane 1, molecular-mass standards; lane 2, soluble proteins (50 μg) collected after concentration on an XM-100 ultrafiltration membrane; lane 3, purified HPPK/DHPS (2 μg) recovered from the folate–agarose column. (B) Immunodetection of the recombinant HPPK/DHPS. Lane 1, soluble proteins (50 μg) collected after concentration on a XM-100 ultrafiltration membrane; lane 2, purified HPPK/DHPS (2 μg); lane 3, molecular-mass standards.

In preliminary experiments, we compared the catalytic properties of the native HPPK/DHPS with those of the recombinant enzyme. The various substrates used for K_m determinations were highly susceptible to degradation and were therefore freshly prepared (see the Materials and methods section). As shown in Table 1, the kinetic parameters of the two types of enzyme were similar, justifying the use of the recombinant protein in this study. However, it must be pointed out that the specific activities reported in Table 1 were measured immediately after the final step of purification. These values decreased with the length of the storage period, suggesting some instability in the purified native and recombinant proteins. This loss of activity was reduced when the enzyme was kept at -80 °C in presence of 10% (v/v) glycerol.

Identification of the dihydropterinPP$_i$ formed by HPPK activity

HPPK activity can be determined in the absence of *p*-ABA either by the amount of [2,8-^{3}H]AMP formed in the presence of [2,8-^{3}H]ATP or the amount of [^{32}P]dihydropterinPP$_i$ formed in the presence of [γ-^{32}P]ATP (see reaction 1). To study the coupling between HPPK and DHPS activities it is necessary to estimate the level of dihydropterinPP$_i$, the product of the first reaction and the substrate for the second. HPLC identification of trace amounts of dihydropterinPP$_i$ is not simple because there is no commercially available labelled standard. Using [γ-^{32}P]ATP or [β,γ-^{32}P]ATP, we identified within our HPLC conditions a peak (retention time, approx. 15 min) that increased with the time course of the reaction. This peak was identified as dihydropterinPP$_i$ for the following reasons. First, it was not present when

Table 1 Kinetic parameters of the HPPK/DHPS protein purified from pea leaf mitochondria and the recombinant HPPK/DHPS protein produced in *E. coli*

The results are means ± S.D. from three to five separate determinations. The specific activities were measured immediately after purification of the enzymes and the K_m values were estimated by direct fitting to Michaelis–Menten curves using non-linear regression and EasyPlot software (Spiral Software).

	Pea HPPK/DHPS		Recombinant HPPK/DHPS	
Kinetic parameter	HPPK + DHPS activity	DHPS activity	HPPK + DHPS activity	DHPS activity
K_m (μM)				
Dihydropterin	1 ± 0.5		2 ± 1	
ATP	70 ± 15		70 ± 15	
DihydropterinPP$_i$		6 ± 4		10 ± 5
p-ABA		1.0 ± 0.5		2.5 ± 1.0
Specific activity (μmol · mg^{-1} · h^{-1})	2.5 ± 0.5	20 ± 4	1.8 ± 0.5	18 ± 4

Table 2 Identification of AMP and dihydropterinPP$_1$ by HPLC

The reaction medium (120 μl) contained 1 μg of the recombinant HPPK/DHPS protein, 100 μM dihydropterin and 200 μM of various labelled ATPs ([2,8-^{3}H]ATP, [α-^{32}P]ATP, [β,γ-^{32}P]ATP and [γ-^{32}P]ATP). After 60 min of incubation, 100 μl of the reaction medium were injected into the HPLC system and the amounts of radioactivity present in the AMP (14 min) and dihydropterinPP$_1$ (15 min) peaks were determined as described in the Material and methods section. The controls (without dihydropterin or without enzyme) were determined with all the various labelled ATP species. The small amount of AMP observed in these controls was present in the initial ATP solution. It was subtracted from all the following data. The values are means $\pm$ S.D. from two determinations for [α-^{32}P]ATP and [β,γ-^{32}P]ATP and at least three determinations for the other conditions; ND, not detectable.

	Specific activity (μmol · mg of protein^{-1})	
	AMP	DihydropterinPP$_1$
No dihydropterin	0.038 $\pm$ 0.008	ND
No enzyme	0.035 $\pm$ 0.009	ND
[^{3}H]ATP	0.53 $\pm$ 0.04	ND
[α-^{32}P]ATP	0.55 $\pm$ 0.03	ND
[β,γ-^{32}P]ATP	ND	0.46 $\pm$ 0.055
[γ-^{32}P]ATP	ND	0.49 $\pm$ 0.04

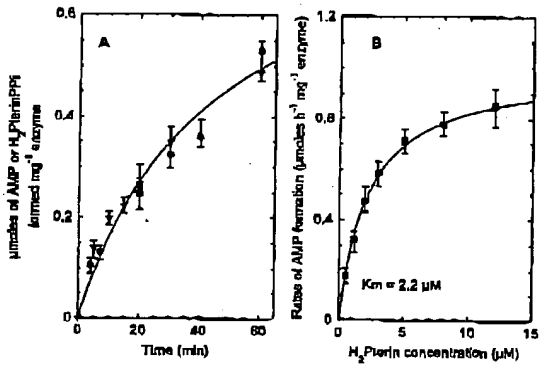


Figure 2 Kinetic study of the HPPK domain

(A) Time-course formation of AMP (▽) and dihydropterinPP$_1$ (H$_2$PterinPP$_1$; ●). HPPK activity is expressed as μmol of products (AMP or dihydropterinPP$_1$) formed/mg of protein. The reaction medium (120 μl) contained 1 μg of recombinant HPPK/DHPS protein, 100 μM dihydropterin and 200 μM ATP. Each point is the mean of two to five measurements. (B) Rate of AMP formation as a function of dihydropterin (H$_2$Pterin) concentration. The reaction medium (120 μl) contained 1 μg of recombinant HPPK/DHPS protein, 200 μM ATP and various concentrations of dihydropterin. The K_m value (2.2 μM) was estimated by direct fitting to the Michaelis–Menten curve using non-linear regression and EasyPlot software (Spiral Software). Each point is the mean from three determinations.

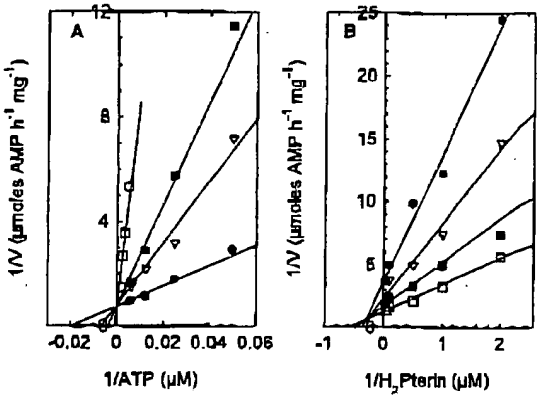


Figure 3 Inhibition of the HPPK activity by dihydropterinPP$_1$

(A) Competitive inhibition by dihydropterinPP$_1$ of ATP in the presence of a saturating concentration of dihydropterin (100 μM). DihydropterinPP$_1$ concentrations were: □, 75 μM; ■, 30 μM; ▽, 7.5 μM; ●, 0 μM. The K_i value was 5 μM. (B) Mixed-type (non-competitive) inhibition by dihydropterinPP$_1$ of dihydropterin (H$_2$Pterin) in the presence of saturating ATP (200 μM). DihydropterinPP$_1$ concentrations were: ●, 30 μM; ▽, 15 μM; ■, 7.5 μM; □, 0 μM. The K_i values were 10 and 15 μM. For both panels each point is the mean from three separate estimations.

Table 3 Inhibition of the HPPK activity by the end products of the reaction

The K_i values for AMP or dihydropterinPP$_1$ versus ATP or dihydropterin were calculated from the initial velocities of the reaction measured in the presence of one particular product (AMP or dihydropterinPP$_1$) and variable concentrations of one particular substrate (ATP or dihydropterin), as described by Segel [28]. The results are means $\pm$ S.D. from three determinations.

	DihydropterinPP$_1$		AMP	
Substrate	K_i (μM)	Type of inhibition	K_i (μM)	Type of inhibition
ATP	5 $\pm$ 3	Competitive	700 $\pm$ 200	Mixed
Dihydropterin	13 $\pm$ 5	Mixed	400 $\pm$ 100	Mixed

either dihydropterin or the enzyme (Table 2) was omitted. Secondly, the peak could not be observed in presence of [2,8-^{3}H]ATP or [α-^{32}P]ATP, a situation where AMP accumulated, indicating that it was not related to the purine ring and did not contain the α-Pi group of ATP. Thirdly, the peak, but not AMP, was observable in the presence of [β,γ-^{32}P]ATP and [γ-^{32}P]ATP (Table 2). The amount of product calculated from the specific activities of the initial ATP substrates was similar in both cases, indicating the presence of two P_i groups. Indeed, if only one P_i was transferred from [β,γ-^{32}P]ATP, the specific activity of the final product would have been half as much. Furthermore, this amount was similar to the amount of AMP formed in the presence of [2,8-^{3}H]ATP or [α-^{32}P]ATP, confirming that the P_i groups were β- and γ-P_i from ATP.

Catalytic properties of the HPPK reaction

In the absence of *p*-ABA and in the presence of saturating concentrations of dihydropterin (100 μM) and ATP (200 μM), the amount of [2,8-^{3}H]AMP and [^{32}P]dihydropterinPP$_1$ increased with time in a stoichiometric manner (Figure 2A). The K_m values for dihydropterin (2.2 $\pm$ 0.7 μM, Figure 2B) and ATP (50 $\pm$ 15 μM, results not shown) were similar to the values reported in Table 1 for the overall reaction (that is, in the presence of *p*-ABA). As shown in Figure 3 and Table 3, dihydropterinPP$_1$ appeared as a competitive inhibitor of ATP (approx. K_i, 5 μM) and a mixed-type (non-competitive) inhibitor of dihydropterin (approx. K_i, 10–15 μM). AMP was a poor inhibitor of the reaction (summarized in Table 3), in agreement with the relatively

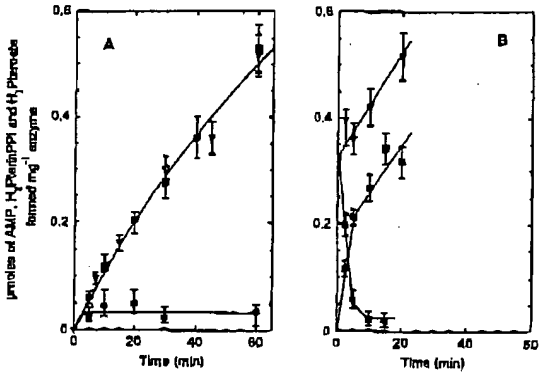


Figure 4 Kinetic study of the coupled HPPK/DHPS reactions

(A) Time-course formation of AMP (▽), dihydropteroate (H_2Pteroate; ■) and dihydropterinPP$_i$ (H_2PterinPP$_i$; ●). The reaction medium (120 μl) contained 1 μg of recombinant HPPK/DHPS, 100 μM dihydropterin, 200 μM ATP and 100 μM p-ABA. (B) Time-course evolution of AMP, dihydropteroate and dihydropterinPP$_i$ in the presence of p-ABA, added 30 min after the addition of the other substrates. The reaction medium (120 μl) contained 1 μg of recombinant HPPK/DHPS, and the initial concentrations of dihydropterin and ATP were 100 and 200 μM, respectively. After 30 min of incubation, 100 μM p-ABA were added and the evolution of the various substrates was monitored as described in the Materials and methods section. Each point is the mean from two to four determinations.

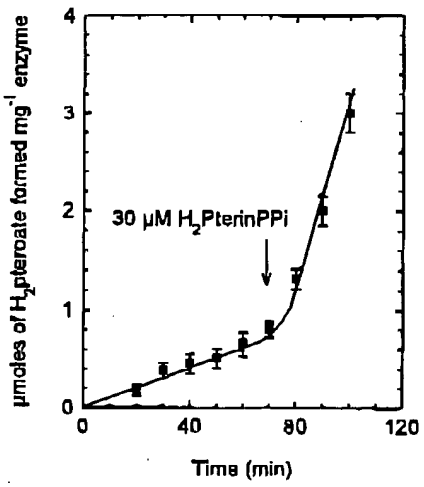


Figure 5 Effect of the addition of dihydropterinPP$_i$ on the coupled activities of HPPK and DHPS

HPPK/DHPS activity is expressed as μmol of dihydropteroate (H_2Pteroate) formed/mg of protein. The reaction medium contained 1 μg of protein, 100 μM dihydropterin, 100 μM p-ABA and 200 μM ATP. Each point is the mean from two determinations. H_2PterinPP$_i$, dihydropterinPP$_i$.

low affinity reported for the monofunctional *E. coli* enzyme [27], and appeared as a mixed-type inhibitor against both ATP and dihydropterin. The crystallographic data obtained with the *E. coli* monofunctional enzyme clearly indicate that both substrates have to be complexed with the enzyme for the one-step reaction of pyrophosphoryl transfer [16], an observation that is indicative of a sequential mechanism [28]. Other studies using analogues of ATP [29] suggest an ordered reaction where ATP is bound first. The product-inhibition pattern is also representative of the type of sequential mechanism (ordered or random) that is involved. The pattern presented in Table 3 is compatible with an ordered sequential Bi Bi mechanism [28] where ATP is bound first and dihydropterinPP$_i$ released last (in a random sequential mechanism in rapid equilibrium, dihydropterinPP$_i$ and AMP would have been competitive inhibitors of both substrates). Thus the plant enzyme mechanism is apparently similar to the one described recently for *E. coli*.

Coupling between HPPK and DHPS activities

The coupling of these two activities was achieved by adding p-ABA (100 μM) together with ATP and dihydropterin. As shown in Figure 4(A), when p-ABA was added at the beginning of the experiment, the rates of AMP and dihydropteroate formation were identical, indicating a good coupling between the two reactions. The velocity of the overall reaction was roughly similar to the one recorded for the HPPK reaction alone (see Figure 2A). In this situation however, dihydropterinPP$_i$ was barely detectable, suggesting either a very rapid transformation through the DHPS activity or limited diffusion in the bulk medium. This last hypothesis is not supported by the accumulation of dihydropterinPP$_i$ observed in Figure 2(A) when the reaction was carried out in the absence of p-ABA. However, a possible conformational change favouring a 'direct' connection between the two catalytic sites (channelling situation) could take place when the two reactions were coupled; that is, when DHPS was working. In order to test this hypothesis, p-ABA was added 30 min after the beginning of the experiment, i.e. when dihydropterinPP$_i$ accumulated in the medium. Indeed, one might expect that a channelling situation would limit the exchange of dihydropterinPP$_i$ between the DHPS catalytic site and the bulk medium. This was apparently not the case since we observed (Figure 4B) a rapid decrease in the dihydropterinPP$_i$ initially present in the external medium, down to the level recorded in Figure 4(A). This drop was concomitant with a symmetrical increase in dihydropteroate. Thus rapid exchanges of dihydropterinPP$_i$ between the DHPS binding site and the external medium were still possible when the two reactions were coupled. The initial rate of dihydropteroate formation shown in Figure 4(B) was about three times higher than the steady-state rate measured in Figure 4(A), suggesting that the DHPS activity could be limited by the availability of dihydropterinPP$_i$. Indeed, as shown in Figure 5, the rate of dihydropteroate synthesis from dihydropterin increased about 9–10-fold when 30 μM dihydropterinPP$_i$ was added to the external medium. Thus the HPPK activity was the rate-limiting step of the overall reaction and the production of dihydropteroate kept pace with that of AMP.

Regulation of the HPPK/DHPS by folate derivatives

In a previous paper [22] we studied some of the kinetic parameters associated with the DHPS activity. This enzyme catalysed a random bireactant reaction and was feedback-inhibited by dihydropteroate, a competitive inhibitor of both dihydropterinPP$_i$ and p-ABA. Interestingly, the monoglutamate forms of dihydrofolate and tetrahydrofolate were also potent inhibitors of the DHPS reaction (Table 4). Like dihydropteroate, these derivatives were competitive with dihydropterinPP$_i$ and p-ABA with similar K_i values for the two substrates. However, the pentaglutamate forms of dihydrofolate and tetrahydrofolate had less effect, and 5-formyl-, 5-methyl- and 5,10-methylene-tetrahydrofolate had no

Table 4 Effect of various folate derivatives on HPPK and DHPS activities

H_2FGlu_1 and H_4FGlu_1 are dihydro- and tetrahydro-folate monoglutamates; H_2FGlu_5 and H_4FGlu_5 are dihydro- and tetrahydro-folate pentaglutamates. These derivatives and dihydropteroate were competitive inhibitors of *p*-ABA and dihydropterinPP$_i$ with similar K_i values for the two substrates. The results are means ± S.D. from three different determinations; NSE, no significant effect for concentrations up to 500 μM.

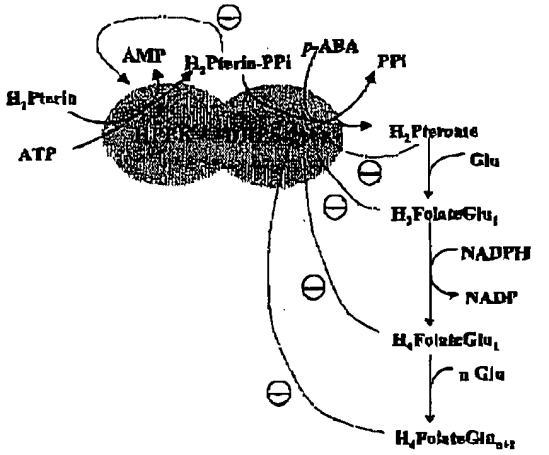
	K_i (μM)	
	HPPK	DHPS
Dihydropteroate	NSE	8 ± 3
H_2FGlu_1	NSE	12 ± 3
H_4FGlu_1	NSE	12 ± 3
H_2FGlu_5	NSE	40 ± 5
H_4FGlu_5	NSE	45 ± 5
5-FormylH_4FGlu_1	NSE	NSE
5-MethylH_4FGlu_1	NSE	NSE
5,10-MethyleneH_4FGlu_1	NSE	NSE

detectable effect on the rate of the DHPS reaction. On the other hand, we were not able to detect any significant change in the velocity of the HPPK reaction in the presence of these various compounds, at least for concentrations up to 500 μM.

DISCUSSION

The plant bifunctional HPPK/DHPS is an important enzyme of the folate biosynthetic pathway. This study represents the first attempt to obtain kinetic information concerning the coupling of the two sequential reactions catalysed by the plant enzyme. The HPPK domain of the protein catalyses the Mg^{2+}-dependent pyrophosphorylation of dihydropterin and our results suggest a sequential Bi Bi ordered system where ATP binds first to the catalytic site and dihydropterinPP$_i$ is released last. These results are in good agreement with the enzyme mechanism described recently for the monofunctional HPPK from *E. coli* [16,29]. Our data also indicate that the level of dihydropterinPP$_i$ remains very low when the two reactions are coupled. Several folate-dependent enzymes involved in sequential reactions were shown to increase their efficiency by channelling intermediates between the catalytic sites. This has been observed with formiminotransferase-cyclodeaminase, an enzyme that presents a channelling specificity that is optimal for the pentaglutamate form of tetrahydrofolate [30]. Another example of channelling of a folate derivative is shown with the bifunctional dihydrofolate reductase/thymidylate synthase. Indeed, the X-ray structure of this enzyme indicates that transfer of dihydrofolate between the two active sites does not occur by a diffusional pathway but rather by electrostatic channelling at the surface of the protein [31]. The term channelling implies that intermediates do not equilibrate with the external medium. In the case of the HPPK/DHPS our results indicate, on the contrary, that exchange of dihydropterinPP$_i$ between the external medium and the protein always remains possible. Although we cannot totally exclude the possibility of a limited diffusion pathway due to the close vicinity of the two catalytic sites, these observations are not in favour of a channelling hypothesis. In our experimental conditions it is likely that the low level of dihydropterinPP$_i$ recorded when the two domains are operating results from the high specific activity of the DHPS versus the HPPK.

In conclusion, it appears from our results that the HPPK activity in plants is the limiting factor of the overall reaction and is highly dependent on the DHPS activity. Indeed, dihydropterinPP$_i$ exerts strong feedback inhibition upon HPPK activity



Scheme 1 Schematic representation of the tetrahydrofolate biosynthetic pathway

The scheme emphasizes the regulation of bifunctional HPPK/DHPS by the different intermediates of the pathway. H_2Pterin, dihydropterin; H_2PterinPP$_i$, dihydropterinPP$_i$; H_2Pteroate, dihydropteroate; H_2FolateGlu, dihydro- ($n = 2$) or tetra- ($n = 4$) hydrofolate glutamate.

and its rate of formation is thus dependent on its utilization by DHPS. As summarized in Scheme 1, the DHPS domain is in turn regulated tightly. Indeed, it is feedback-inhibited by dihydropteroate, the product of the reaction [22], and it is also inhibited strongly by the monoglutamate forms of dihydrofolate and tetrahydrofolate, two intermediates synthesized downstream in the folate biosynthetic pathway. From all these data it is tempting to postulate that the DHPS domain of the HPPK/DHPS protein is a key regulatory point of the folate biosynthetic pathway in plants. If this holds true, this mitochondrial enzyme could be a good target for folate enhancement in plants of nutritional interest. Indeed, serious diseases are correlated with folate deficiency and one possible answer to these health problems could be to increase the folate content of plant food (for a review, see [32]). We are currently investigating the effects of varying the level of expression of the bifunctional HPPK/DHPS on the folate status in plants.

We gratefully thank Professor D. Macherel (University of Angers, Angers, France) for helpful discussions.

REFERENCES

1 Rébeillé, F. and Douce, R. (1999) Folate synthesis and compartmentation in higher plants. In Regulation of Primary Metabolic Pathways in Plants (Kruger, N. J., Hill, S. A. and Ratcliffe, R. G., eds.), pp. 53–99, Kluwer Academic Publishers, Dordrecht
2 Shiota, T. (1994) Biosynthesis of folate from pterin precursors. In Folates and Pterins, vol. 1 (Blakley, R. L. and Benkovic, S. J., eds.), pp. 121–134, Wiley Interscience, New York
3 Hong, Y.-L., Hossler, P. A., Calhoun, D. H. and Meshnick, S. R. (1995) Inhibition of recombinant *Pneumocystis carinii* dihydropteroate synthetase by sulfa drugs. J. Antimicrob. Chemother. 35, 1756–1763
4 Lopez, P., Espinosa, M., Greenberg, B. and Lacks, S. A. (1987) Sulfonamide resistance in *Streptococcus pneumoniae*: DNA sequence of the gene encoding dihydropteroate synthase and characterization of the enzyme. J. Bacteriol. 169, 4320–4326
5 Zhang, Y. and Meshnick, S. R. (1991) Inhibition of *Plasmodium falciparum* dihydropteroate synthetase and growth *in vitro* by sulfa drugs. Antimicrob. Agents Chemother. 35, 267–271

6 Brooks, D. R., Wang, P., Read, M., Watkins, W. M., Sims, P. F. G. and Hyde, J. E. (1994) Sequence variation of the hydroxymethyldihydropterin pyrophosphokinase: dihydropteroate synthase gene in lines of the human malaria parasite, *Plasmodium falciparum*, with differing resistance to sulfadoxine. Eur. J. Biochem. **224**, 397–405

7 Fermer, C. and Swedberg, G. (1997) Adaptation to sulfonamide resistance in *Neisseria meningitidis* may have required compensatory changes to retain enzyme function: kinetic analysis of dihydropteroate synthases from *N. meningitidis* expressed in a knockout mutant of *Escherichia coli*. J. Bacteriol. **179**, 831–837

8 Triglia, T., Menting, J. G. T., Wilson, C. and Cowman, A. F. (1997) Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in *Plasmodium falciparum*. Proc. Natl. Acad. Sci. U.S.A. **94**, 13944–13949

9 Dallas, W. S., Gowen, J., Ray, P. H., Cox, J. and Dev, I. K. (1992) Cloning, sequencing, and enhanced expression of the dihydropteroate synthase gene of *Escherichia coli* MC4100. J. Bacteriol. **174**, 5961–5970

10 Kellam, P., Dallas, W., Ballantine, S. P. and Delves, C. J. (1995) Functional cloning of the dihydropteroate synthase gene of *Staphylococcus haemolyticus*. FEMS Microbiol. Lett. **134**, 165–169

11 Achari, A., Somers, D. O., Champness, J. N., Bryant, P. K., Rosemond, J. and Stammers, D. K. (1997) Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. Nat. Struct. Biol. **4**, 490–497

12 Hampele, I. C., D'Arcy, A., Dale, G. E., Kostrewa, D., Nielsen, J., Oefner, C., Page, M. G. P., Schönfeld, H.-J., Stüber, D. and Then, R. L. (1997) Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. J. Mol. Biol. **268**, 21–30

13 Talarico, T. L., Ray, P. H., Dev, I. K., Merill, B. and Dallas, W. S. (1992) Cloning, sequence analysis, and overexpression of *Escherichia coli* folK, the gene coding for 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase. J. Bacteriol. **174**, 5971–5977

14 Lopez, P. and Lacks, S. A. (1993) A bifunctional protein in the folate biosynthetic pathway of *Streptococcus pneumoniae* with dihydropterin aldolase and hydroxymethyldihydropterin pyrophosphokinase activities. J. Bacteriol. **175**, 2214–2220

15 Stammers, D. K., Achari, A., Somers, D. O., Bryant, P. K., Rosemond, J., Scott, D. L. and Champness, J. N. (1999) 2.0 Å X-ray structure of the ternary complex of 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase from *Escherichia coli* with ATP and a substrate analogue. FEBS Lett. **456**, 49–53

16 Blaszczyk, J., Shi, G., Yan, H. and Ji, X. (2000) Catalytic center assembly of HPPK as revealed by the crystal structure of a ternary complex at 1.25 Å resolution. Structure, **8**, 1049–1058

17 Hennig, M., Dale, G. E., D'Arcy, A., Danel, F., Fischer, S., Gray, C. P., Jolidon, S., Müller, F., Page, M. G. P., Pattison, P. and Oefner, C. (1999) The structure and function of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Haemophilus influenzae*. J. Mol. Biol. **287**, 211–219

18 Triglia, T. and Cowman, A. F. (1994) Primary structure and expression of the dihydropteroate synthase gene of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. U.S.A. **91**, 7149–7153

19 Allegra, C. J., Boarman, D., Kovacs, J. A., Morrison, P., Beaver, J., Chabner, B. A. and Masur, H. (1990) Primary structure and expression of the dihydropteroate synthase gene of *Plasmodium falciparum*. J. Clin. Invest. **85**, 371–379

20 Volpe, F., Ballantine, S. P. and Delves, C. J. (1993) The multifunctional folic acid synthesis *fas* gene of *Pneumocystis carinii* encodes dihydroneopterin aldolase, hydroxymethyldihydropterin pyrophosphokinase and dihydropteroate synthase. Eur. J. Biochem. **218**, 449–458

21 Neuburger, M., Rébeillé, F., Jourdain, A., Nakamura, S. and Douce, R. (1996) Mitochondria are a major site for folate and thymidylate synthesis in plants. J. Biol. Chem. **271**, 9466–9472

22 Rébeillé, F., Macherel, D., Mouillon, J. M., Garin, J. and Douce, R. (1997) Folate biosynthesis in higher plants: purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. EMBO J. **16**, 947–957

23 Scrimgeour, K. G. (1980) Methods for reduction, stabilization, and analyses of folates. Methods Enzymol. **66**, 517–523

24 Temple, Jr, C. and Montgomery, J. A. (1984) Chemical and physical properties of folic acid and reduced derivatives. In Folates and Pterins, vol. 1 (Blakley, R. L. and Benkovic, S. J., eds.), pp. 61–120, Wiley Interscience, New York

25 Pfleiderer, W. (1984) Chemistry of naturally occuring pterins. In Folates, Pterins, vol. 2 (Blakley, R. L. and Benkovic, S. J., eds.), pp. 43–114, Wiley Interscience, New York

26 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**, 680–685

27 Shi, G., Gong, Y., Savchenko, A., Zelkus, J. G., Xiao, B., Ji, X. and Yan, H. (2000) Dissecting the nucleotide binding properties of *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase with fluorescent 3′(2)′-o-anthraniloyladenosine 5′-triphosphate. Biochim. Biophys. Acta **1478**, 289–299

28 Segel, I. H. (1975) Enzyme Kinetics, John Wiley and Sons, New York

29 Bermingham, A., Bottomley, J. R., Primrose, W. U. and Derrick, J. P. (2000) Equilibrium and kinetic studies of substrate binding to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Escherichia coli*. J. Biol. Chem. **275**, 17962–17967

30 Paquin, J., Baugh, C. M. and MacKenzie, R. E. (1985) Channeling between the active sites of formiminotransferase-cyclodeaminase. Binding and kinetic studies. J. Biol. Chem. **260**, 14925–14931

31 Knighton, D. R., Kan, C. C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M. and Matthews, D. A. (1994) Structure of and kinetic channelling in bifunctional dihydrofolate reductase-thymidylate synthase. Nat. Struct. Biol. **1**, 186–194

32 Scott, J., Rébeillé, F. and Fletcher, J. (2000) Folic acid and folates: the feasibility for nutritional enhancement in plant foods. J. Sci. Food Agric. **80**, 795–824

Received 13 September 2001/12 November 2001; accepted 31 January 2002